91602
new 82-10

LEGAL NOTICE

This report was prepared by Ontario Hydro and, while effort has been made to assure the accuracy and applicability of the material contained herein, neither Ontario Hydro nor any person acting on its behalf:

(a) makes any warranty, express or implied with respect to the use of any information, apparatus, method or process disclosed in this report or that such use may not infringe privately owned rights; or

(b) assumes any liabilities with respect to the use of, or for damages resulting from the use of, any information, apparatus, method or process disclosed in this report.

COMPLIMENTARY COPY
Please Credit Ontario Hydro Research if Reproduced in Whole or in Part

ontario hydro
research division

ESTIMATION PROCEDURES FOR FRACTURE PARAMETER DISTRIBUTIONS

Report No. 87-277-K

A.H.C. Chan and A.L. Punhani
Reliability & Statistics Section
Operations Research Department

ABSTRACT

In evaluating a plutonic rock mass as a permanent nuclear waste repository host site, it is necessary to predict the groundwater flow through the fracture network in the rock. Consequently, this network needs to be determined. However, data on fractures are sparse and expensive to collect. Under the circumstances, the Monte Carlo simulation method was used for the purpose. This method requires the probability distributions of the fracture parameters. In this report, procedures are devised to estimate these probability distributions. Computer codes for these procedures are provided and the procedures are illustrated by examples using data from the Atikokan test site and the Underground Research Laboratory.

Work done for
Atomic Energy of Canada Limited
Whiteshell Nuclear Research Establishment
under the Technical Assistance Program

job	file	date	report no.
72-126-461	809.730	January 6, 1988	87-277-K

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	1
2.	ESTIMATION OF THE FRACTURE PARAMETER DISTRIBUTION	2
	2.1 Parametric Representation of Fractures	2
	2.2 Outliers of the Estimation Procedures	4
	2.3 Determination of the Joint Sets	5
	2.3.1 The Measure of Proximity	5
	2.3.2 The Clustering Algorithm	6
	2.3.3 Criterion for Determining the Optimal Number of Clusters	6
	2.4 Distribution of the Orientation	8
	2.4.1 The Spherical Beta Distribution	8
	2.4.2 The Goodness-of-Fit Test	9
	2.5 Distribution of Fracture Dimensions	9
	2.6 Distribution of Density	9
	2.7 Extrapolation of the Estimated Distribution	10
3.	AN EXAMPLE: THE ATIKOKAN DATA	11
	3.1 Determination of the Joint Sets	11
	3.2 Estimates of the Orientation Distributions	13
	3.3 Estimates of the Density Distribution	15
4.	SUMMARY OF DISCUSSIONS	16
REFERENCES		18
APPENDIX I -	Source Code for the Subroutine KMNS	19
APPENDIX II -	Source Code for the Subroutine BETPLT	23
APPENDIX III -	The Kolmogorov - Smirnov - Durbin Statistics	28
APPENDIX IV -	Estimation of Fracture Dimensions Distribution from the Distribution of Surface Fracture Lengths	30
APPENDIX V -	Results From the URL Borehole Data	33
APPENDIX VI -	Results From the URL Excavation Data	37
DISTRIBUTION		Last Page

87-277

91601
new 82-10

ontario hydro
research division

To Dr. D. Mills
Director of Research

ESTIMATION PROCEDURES FOR FRACTURE PARAMETER DISTRIBUTIONS

1. INTRODUCTION

The most favoured method for permanent nuclear waste disposal in Canada is to isolate the radioactive waste in a repository located deep in a stable plutonic rock mass. One of the most important selection criteria for a repository host rock is the inertness of its groundwater. Fractures exist inside plutonic rock masses and groundwater flows through the network formed by these fractures. Hence, in assessing the suitability of a potential host rock for nuclear waste disposal, the groundwater flow through the fracture networks in the rock mass needs to be determined.

A numerical method to calculate the discrete network flow (FLOWNET computer code [1]) was developed at the Whiteshell Nuclear Research Establishment of AECL. However, the FLOWNET procedure requires complete specifications of a fracture network for its input. In practice, only limited information is available on the network obtained from expensive and widely separated boreholes. Thus, a flexible method to estimate the fracture network is needed.

In [2], a parametric representation of fracture networks developed in Ontario Hydro Research Division was described. This representation is based on the assumption of a joint probability distribution for the fracture parameters. If this probability distribution is given, the fracture network can be estimated by the Monte Carlo simulation technique. Thus, the problem becomes that of estimating the joint probability distribution of the fracture parameters. One of the main objectives of this report is to develop procedures for estimating such a distribution along with computer codes to implement these procedures. This is discussed in Section 2.

The final objective of this report is to evaluate these procedures by applying them to the Atikokan and the Underground Research Laboratory (URL) borehole data. Results for the Atikokan data are given in Section 3. In Appendices V and VI, similar results for the URL borehole and excavation data are given. The computer codes KMNS and BETPLT which are used to implement these procedures are given in Appendices I and II. In Appendix III, the rationale behind the development of the computer code BETPLT is discussed.

An added complication in the estimation procedure is the fact that the samples from the boreholes may not be representative. Fracture samples are usually obtained from a relatively small portion (testing area) of the rock mass; and only if the fractures are distributed homogeneously, can the samples be considered representative. But the assumption of homogeneity is unlikely to be true. If it is not, the estimated distribution is not applicable to the entire rock mass. Hence, some criteria for extrapolating the estimated distribution to the

87-277

entire rock are needed. This will be the subject matter of further studies.

To recapitulate, the following tasks were carried out:

1. The development of procedures for determining the joint probability distribution of the fracture parameters.

2. The development of computer codes to implement these procedures.

3. The evaluation of these procedures with available data from the Atikokan and the URL sites.

2. ESTIMATION OF THE FRACTURE PARAMETER DISTRIBUTION

Consider a rock mass as a region in a Cartesian coordinate system. Then a fracture can be considered as a small volume inside this rock mass (Figure 1). Although fractures may have almost unlimited variety of shapes, it is impractical to attempt to mathematically model all these possibilities. Therefore, a tractable model for fractures must assume that fractures take on only a finite number of regular geometric forms. These geometric forms must be completely determined by some measurable parameters of the fractures (ie these geometric forms, and hence the fractures, must be represented by a relatively small number of fracture parameters).

2.1 Parametric Representation of Fractures

The only measurable geometric attributes of a fracture are its orientation (strike and dip angles), its location (the coordinates at which the fracture intersect the borehole) and, possibly, its dimensions (span and aperture). The simplest geometric solid in a Cartesian coordinate system that can be defined by these attributes is either a disk or a prism. In [2], a fracture is idealized as a prism. However, a circular representation can provide a method to estimate the distribution of fracture dimensions. Hence, in this report, fractures are represented by circular disks. The combinatorial procedures developed in [3] to calculate the fracture network can be easily modified to accommodate this representation. Figure 2 shows that an idealized fracture is represented by its location, orientation and its dimensions.

FIGURE 1

A FRACTURE IN A PLUTONIC ROCK MASS

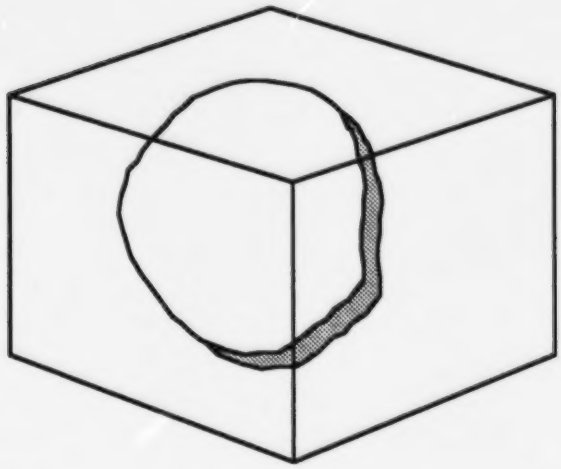
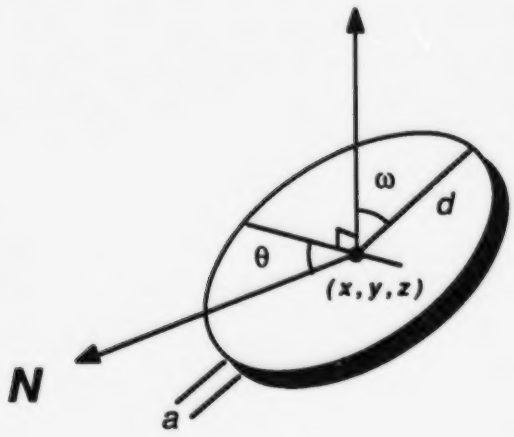


FIGURE 2

AN IDEALIZED FRACTURE



The parameters are defined as follows:

The orientation parameters:

θ = the strike angle of a fracture;
ω = the dip angle of a fracture.

The location parameters:

(x,y,z) = the Cartesian coordinates of the centre of the disk.

The dimension parameters:

d = the diameter of the disk;
a = the aperture of a fracture.

The parameters for the location reference points are simply indications of the fracture density, κ, defined as the number of fractures per unit volume. However, this parameter is not observable. But from the borehole data, the spacing between fractures is observable. The reciprocal of spacing is the linear fracture density, ρ, defined as:

$$\rho = M / L$$

where M is the number of fractures meeting a vertical line (borehole) and L is the length of the vertical line. Together with ρ and the fracture dimension, the fracture density κ can be determined. In [3], a procedure was provided to generate location reference points from a given density distribution. Hence, in this study, the fracture density is considered instead of (x,y,z). Consequently, the joint probability distribution of the parameters (ρ,d,a,θ,ω) needs to be estimated.

2.2 Outlines of the Estimation Procedures

Let $P(\rho,d,a,\theta,\omega)$ denote the joint probability distribution of the parameters (ρ,d,a,θ,ω). Such a distribution can be decomposed into a product of conditional distributions [4]. In particular, a possible decomposition of P is

$$P(\rho,d,a,\theta,\omega) = F(d,a|\rho,\theta,\omega)G(\rho|\theta,\omega)H(\theta,\omega) \qquad (1)$$

where F and G are conditional distributions. Thus, estimating P is equivalent to estimating F, G and H. The estimation procedure can be outlined by the following tasks:

Task 1. Determination of the joint sets.
Task 2. Estimation of the orientation distribution $H(\theta,\omega)$ of each joint set.
Task 3. Estimation of the dimension distribution $F(d,a)$ of each joint set.
Task 4. Estimation of the density distribution $G(\rho)$ of each joint set.

As discussed in the Introduction, the estimated distribution must be extrapolated to the entire

rock mass. So, the final step of the estimation procedure is,

Task 5. Extrapolation of the estimated distribution to the entire rock mass.

Remark: Because of insufficient relevant data, procedures for task 5 were not developed in this report. However, they are briefly discussed in section 2.7.

2.3 Determination of the Joint Sets

A joint set is defined to be a group of fractures having similar orientations. Therefore, to determine the joint sets, a Statistical Cluster Analyses technique was used on the collection of all fracture orientations. There are three initial specifications which must be made in order to implement the Cluster Analysis. They are:

1. Definition of a measure of proximity;
2. Selection of an appropriate algorithm for this measure of proximity;
3. Adoption of a criterion for determining the optimal number of clusters.

2.3.1 The Measure of Proximity

The measure of proximity must indicate the similarity between orientations of two fractures. The best indicator is the angle between the normal vectors of the planes of two fractures (ie, 0 and 180 degrees being the most similar and 90 degrees the least). This angle can be defined as follows:

Let (θ,ω) and (Θ,Ω) denote the orientations of two fractures f and F. Their normal vectors are

$$\mathbf{u} = (u_1,u_2,u_3)$$

and

$$\mathbf{U} = (U_1,U_2,U_3)$$

respectively, where

$$\begin{aligned} u_1 &= \cos\theta \sin\omega \\ u_2 &= -\sin\theta \sin\omega \\ u_3 &= \cos\omega \\ U_1 &= \cos\Theta \sin\Omega \\ U_2 &= -\sin\Theta \sin\Omega \\ U_3 &= \cos\Omega \end{aligned} \qquad (2)$$

The angle between **u** and **U** is defined by

$$\delta(\mathbf{u},\mathbf{U}) = \arccos(u_1U_1 + u_2U_2 + u_3U_3) \qquad (3)$$

Now, this angle takes on values between 0 and 180 degrees. But supplementary angles represent the same difference in orientation. Hence, by restricting it to acute angles, a measure of proximity can be defined by

$$\Delta(\mathbf{u},\mathbf{U}) = \min \{\delta(\mathbf{u},\mathbf{U}),\delta(-\mathbf{u},\mathbf{U})\} \qquad (4)$$

This way, Δ only takes on values between 0 and 90 degrees and provides a good indication of the similarity between two sets of orientations.

This measure, however, is difficult to work with in Cluster Analysis. Fortunately, it is equivalent to the Euclidean distance which is more suitable for clustering algorithms. To see this, consider a normal vector as a point on the unit sphere, ie, the sphere centered at the origin with radius equal to unity. Then, the Euclidean distance $D(\mathbf{u},\mathbf{U})$ of $\mathbf{u}$ and $\mathbf{U}$ has the following relationship with $\Delta(\mathbf{u},\mathbf{U})$:

$$D(\mathbf{u},\mathbf{U}) = 2\sin(\Delta(\mathbf{u},\mathbf{U})/2)$$

This means that D is equivalent to Δ. The function Δ will be defined as the measure of proximity because it is more suitable for the clustering algorithms.

2.3.2 The Clustering Algorithm

There are many possible clustering algorithms. An algorithm called the K-Mean clustering [6] was selected based on the efficiency of its modified version [7]. A FORTRAN subroutine KMNS is listed in Appendix I.

For input, the KMNS code requires a list of all orientations in the form of normal vectors (t_i,u_i,v_i), $i=1,2,3,...,n$; and the number of clusters, K, to be grouped. In output, it provides

1. A list of the members in each cluster;
2. The coordinates of the centers of K clusters;
3. The within cluster sum of squares of each cluster.

Using the terminology of [5], a cluster is called a joint set.

2.3.3 Criterion for Determining the Optimal Number of Clusters

Since the number of clusters (ie joint sets) is an input at each application of this algorithm, a criterion for selecting an optimal number must be adopted. There are two possible ways to make this decision. Both involve subjective judgments. First, by inspecting the stereoplot of the fracture orientation visually [5], a reasonable subjective decision can be made on the number of joint sets.

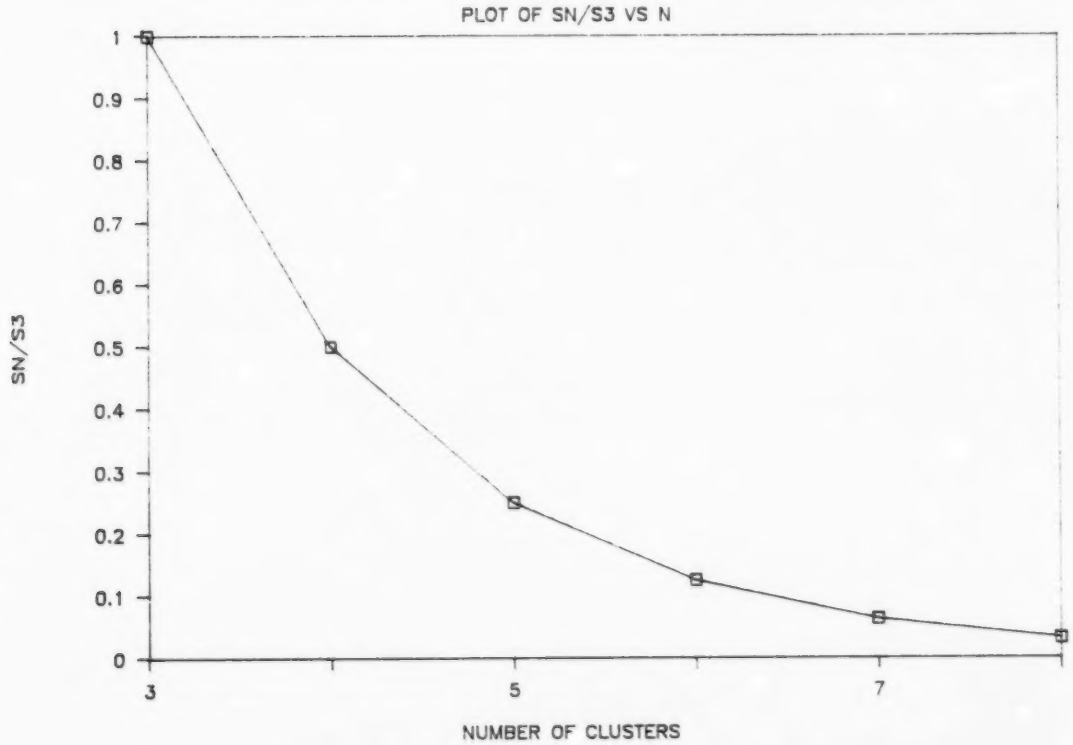
Alternatively, the following procedure can be used:

Let N be the number of joint sets to be selected. Set $N = 3$ (the K-Mean algorithm requires $N > 2$). Let S_N denote the total sum of squares of all the N clusters.

1. Apply KMNS to get S_N;
2. Increase N by 1 and apply KMNS to get S_{N+1};
3. Repeat Step 2 until a sequence $S_3, S_4, S_5, \ldots$ is obtained (This series should have about 5 to 6 terms);
4. Plot S_N/S_3 versus N (see Figure 3);
5. Choose K to be the number such that S_K/S_3 is suitably small (eg 0.3, 0.4 or 0.5).

Remark: A criterion developed for use in regression analysis, called the Mallow's Cp-statistic [8] could be adopted to find a minimum in the sum of squares versus number of fitted parameters (clusters) plot. However, this would still require some arbitrary selection of "residual" sum of squares.

FIGURE 3



2.4 Distribution of the Orientation

In the above section, the orientation of a fracture was transformed to a point on the unit sphere. Therefore, the distribution of orientation can be considered as a spherical distribution. To estimate this distribution, the goodness-of-fit test of the data against a number of possible distributions was tried. The best fitted one will be considered as the distribution for the orientation. It must be pointed out that having passed the goodness-of-fit test is not a definitive proof for the distribution of the orientations. However, it provides a good enough approximation for the simulation purpose.

Therefore, it is desirable to find a distribution which is robust enough to fit the data from different types of populations. Although a robust distribution may not be the exact one for the orientation, it will provide a good approximation to the true distribution.

Finally, if T is the transformation defined by

$$T(\theta,\omega) = \mathbf{u} = (u_1,u_2,u_3)$$

where the u_i's are defined in (2), and $\Psi(\mathbf{u})$ is the estimated spherical distribution of $\underline{u}$, then

$$H(\theta,\omega) = \Psi[T(\theta,\omega)]$$

is the distribution for (θ,ω). In the following, a special spherical distribution and a goodness-of-fit test which are needed in the estimation procedures are discussed.

2.4.1 The Spherical Beta Distribution

The fracture orientations are bounded, since they are less than 180 degrees. If this is a one dimension variable, the Beta distribution will be one of the most robust one [9]. Although, there is no spherical analogue of the Beta distribution, it can be defined in the following manner:

Definition:

> Let $\mathbf{u}=(u_1,u_2,u_3)$ be a random variable on the unit sphere. The variable $\underline{u}$ is said to have a Spherical Beta distribution around a center $\mathbf{v}$ if its probability density function has the form
>
> $$\psi(\mathbf{u}) = B(\alpha,\beta)^{-1}(\Delta(\mathbf{u},\mathbf{v})/90)^{\alpha-1}(1 - \Delta(\mathbf{u},\mathbf{v})/90)^{\beta-1} \qquad (5)$$
>
> where $\alpha > 1$, $\beta > 2$, $\Delta(\mathbf{u},\mathbf{v})$ is defined by (4) in Section 2.3 and $B(\alpha,\beta)$ is the Beta function defined by
>
> $$B(\alpha,\beta) = \int_0^1 \zeta^{\alpha-1}(1 - \zeta)^{\beta-1}d\zeta$$

The KMNS code provides the coordinates of the centers for each joint set and the list of all its members. Then, these data will be fitted to a Spherical Beta distribution.

2.4.2 The Goodness-of-Fit Test

Suppose that $\mathbf{v}$ is the center of a joint set (cluster) and $\mathbf{u}_1,\mathbf{u}_2,...,\mathbf{u}_n$ are its members. Then compute

$$x_i = \Delta(\mathbf{u}_i,\mathbf{v})/90$$

where Δ is defined in (4). To determine if $\mathbf{u}_1,\mathbf{u}_2,...,\mathbf{u}_n$ are from a Spherical Beta population, it is equivalent to determining if the x_is are from a one dimensional Beta population.

A FORTRAN subroutine BETPLT is developed to fit the data to a selected distribution (amongst Beta, Gamma and Exponential) as well as testing the goodness-of-fit. The listing of BETPLT is given in Appendix II. The Kolmogorov-Smirnov-Durbin procedure was used for the goodness-of-fit test. The mathematical concept of this subroutine is given in Appendix III.

2.5 Distribution of Fracture Dimensions

There is no aperture data available to Ontario Hydro when this report was prepared. If the aperture is assumed to be independent from all other parameters, then, without lost of generality, it can be assumed to be constant. Therefore, it suffices to consider only the distribution of the diameters. If $\Psi(\gamma)$ is the probability distribution function of the surface lengths and $H(\psi)$ is the distribution of the dip angles, then the distribution $F(d)$ of diameters is given by the equation:

$$F(y) = 1 - C \int_{2y}^{\infty} [(\gamma/2)^2 - y^2]^{-1/2} \, d\Psi(\gamma) \qquad (6)$$

where $C = Z_o/[\pi \int_\Omega \sin\omega dH(\omega)]$, Z_o is the depth of the rock mass, and $\Omega = [0,\pi/2]$. The proof of this result is given in Appendix IV. The concept of this methodology is given in [10]. Due to lack of data, this procedure was not tested and verified.

2.6 Distribution of Density

As discussed in Section 2.1, the density parameter is assumed to be independent of the orientation within each joint set. Hence, it suffices to estimate the marginal distribution of the density within each joint set. Now, density is defined as:

$$\rho = \text{number of fractures per meter}$$

To determine the density distribution from the data, the following procedure is used.

Let (x_r,y_r,z_r), $r = 1,2,...N$, be the observed locations of the fractures of the joint set within a borehole (as provided by the KMNS code). Let $\mathbf{v} = (v_1,v_2,v_3)$ be its center. Let m be a predetermined length, eg $m = 25$ meters. Finally, set $k = 0$.

1. Compute

$$p_r = v_1 x_r + v_2 y_r + v_3 z_r.$$

2. Sort p_r's so that

$$p_1 < p_2 < \ldots < p_n.$$

3. Define a sequence $\{ s_k \}$ as follows:

   (a) Set k to $k+1$, $i=1$ and $j=0$.

   (b) Determine

$$n(k) = \text{number of } p_r\text{'s between } mi \text{ and } mj \qquad (7)$$

   where mi and mj are multiples of i and j by m. If $n(k)=0$, then set i to $i+1$ and determine $n(k)$ by (7). Repeat this until $n(k) > 0$ or a fixed number, say, $n(k) > 2$.

   Define

$$s_k = n(k)/(i-j)m \qquad (8)$$

   (c) Set k to $k+1$, i to $i+1$ and j to i. Then, go to (b) to find s_k. Repeat this until all p_r's have been enumerated.

4. Repeat steps 1 to 3 for each borehole.

After s_k's from all boreholes were obtained, fit the data to a robust distribution using the subroutine BETPLT. The one that fits best will be the desired distribution $G(\rho)$. The most likely candidate is the Gamma distribution.

Remark: In [2], the parameter studied was the fracture spacing rather than of density. However, spacing is the reciprocal of density. Hence, the network generating algorithm presented in [2] will not be affected. The URL borehole data supplied by GSC do not have the location parameters. The determination of the density is slightly different (see Appendix V for details).

2.7 Extrapolation of the Estimated Distribution

In the Introduction, the need to devise procedures for extrapolating the estimated distribution was discussed. These procedures must be based on some relationships between fractures in rock masses and other geological phenomena. One such possibility could be the relationship between the surface and the subsurface fractures. Since surface fractures are observable without borehole drilling, such data will be easier to obtain.

Unfortunately, up to this point, there is no consensus among the geologists on whether such a relationship can be established or it even exists. Therefore, to undertake any such task is

purely academic. But, as a first step toward realizing such a relationship, there are at least two exercises that can be performed which may be useful. They are:

1. Determine the joint sets of the surface fractures. If the joint sets of the surface fracture are vastly different from those of the subsurface fractures, then no relationship can be established. On the other hand, if they are in substantial agreement, then a functional relationship between depth and density can be established.

2. Establish a relationship between density and depth of each of the joint sets.

3. AN EXAMPLE: THE ATIKOKAN DATA

To evaluate the procedures described in the above section, the Geological Survey of Canada (GSC) provided Ontario Hydro with the Atikokan borehole data. The format of the data, which is stored as an ASCII file, is as follows:

Each record consists of eleven fields pertaining to a fracture. They are

(X,Y,Z,FT,SK,DP,TF,FF,BN,LB,AV)

with format

(3F10.3,A2,F3.0,F2.0,F6.1,A10,I2,F8.3,F4.1)

where

X,Y,Z = the coordinates of the intersection of the fracture and the borehole,
FT = fracture type,
SK = strike angle,
DP = dip angle,
TF = thickness of the fracture,
FF = fracture filling type,
BN = borehole number,
LB = length along the borehole,
AV = orientation adjustment value.

Remark: Similar examples for the URL borehole and excavation data are given in the Appendices V and VI, respectively.

3.1 Determination of the Joint Sets

Fractures whose orientation was not measurable were excluded from the Cluster Analysis. The orientation data were transformed to their normal vectors before applying the clustering

routine. Also it was decided that the optimal number of clusters is six in accordance with [5]. However, in Appendices V and VI, the number of joint sets for the URL data were determined using the method outlined in Section 2.3.3.

The following tables summarize the results of the cluster analysis. Table 1 provides the average orientation of each of the clusters, while Table 2 compares the results obtained from the cluster analysis and that from a study by the GSC [5].

TABLE 1

Cluster Analysis Results on the Orientation

Cluster	Center of Orntns		Within Clsrs	No. of
	Strike	Dip	Sum of Squrs	Obsvns
1	358.8	9.7	290.4	2583
2	297.5	65.4	293.6	2113
3	54.1	47.8	255.9	1427
4	213.1	44.4	282.1	1802
5	129.0	54.5	187.1	1109
6	350.0	76.3	361.9	2559

Remark: The above results were obtained from the original data file provided by the GSC which does not include the value AV. This is rather important because the AV value will give more weight to the fractures as they are almost parallel to the angles of the boreholes. This may explain why the results differ somewhat from the GSC results as outlined in Table 2.

TABLE 2

Comparison Between the GSC Results and the Cluster Analysis Results

GSC		Cluster Analysis	
Jt. Set	Orientation	Cluster	Orientation
1	335/67.5	6	350.0/76.3
2	300/67.5	2	297.5/65.4
3	213/40.0	4	213.0/44.4
4	60/32.0	3	54.1/47.8
5	352/13.6	1	358.8/ 9.8
6	8/75.0	6	350.0/76.3

Remark: Cluster 5 does not match any joint set. This may be explained by the results of hypothesis testing in Table 3. But cluster 6, on the other hand, matches both joint sets 1 and 6. This may be due to the fact that the two orientations are very similar and that cluster 6 is situated between the two joint sets.

3.2 Estimates of the Orientation Distributions

From the cluster analysis, the cluster numbers JS are added to the data file. So, from the data file, all the fractures from a cluster (joint set) can be selected. As described in Section 2.3, the Spherical Beta distribution is probably the most robust one for the data. The results are summarized in Table 3.

TABLE 3

Results of Fitting the Orientation to the Spherical Beta Distributions

Cluster	α	β	P_o	H_o
1	5.35	25.66	0.94	Accepted
2	2.75	12.30	0.90	Accepted
3	2.75	16.63	0.99	Rejected
4	3.92	12.65	0.89	Accepted
5	0.77	5.15	1.00	Rejected
6	4.20	15.12	0.78	Accepted

where α and α are the estimated parameters of the Spherical Beta distribution (see Definition in 2.4.1); H_o is the null hypothesis that the orientation of a cluster have a Spherical Beta distribution; P_o is the probability to reject H_o.

Accepting the hypothesis H_o means that the empirical distribution formed by the data is not distinguishable from a Spherical Beta distribution based on a 1% significance level. Hence, Spherical Beta distributions provide good approximations of the distribution of orientations for clusters 1, 2, 4 and 6. For example, Figure 4 shows the plot of the empirical distribution versus the fitted Spherical Beta distribution for cluster 1. Clearly, they are very similar.

On the other hand, rejecting H_o does not mean the Spherical Beta does not approximate the empirical distribution. For example, Figure 5 shows that the empirical distribution of the orientation of cluster 3 is very close to a Spherical Beta distribution. Although it was rejected, for the purpose of simulation, cluster 3 may also be considered to have a Spherical Beta distribution.

Figure 6 suggests that no continuous distribution will fit the data in cluster 5 due to the large weight at 5 degree. This may also explain why this cluster does not match any joint set given in [5].

FIGURE 4

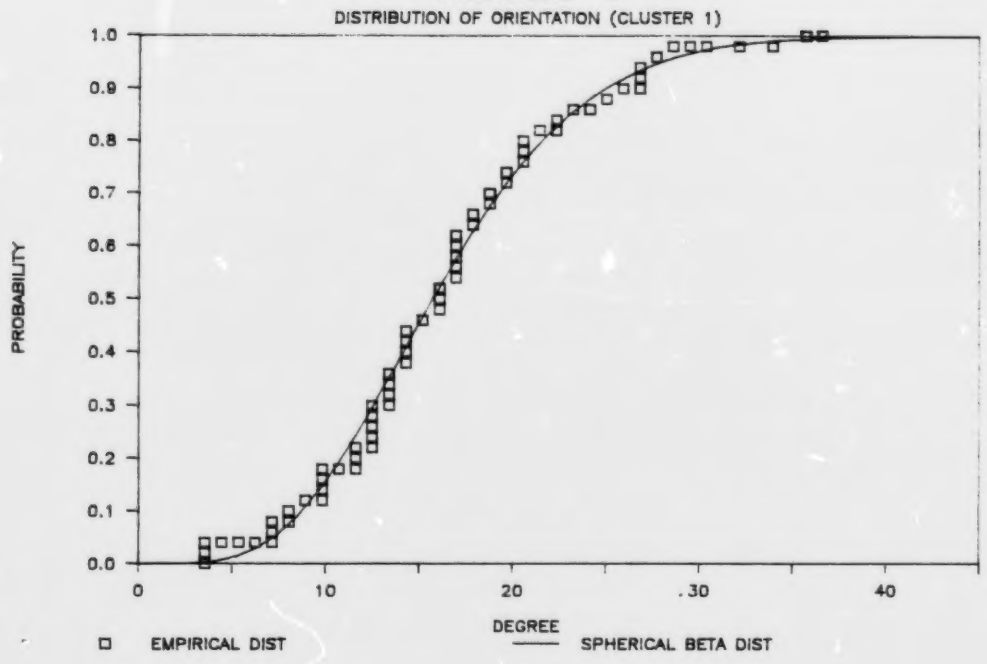


FIGURE 5

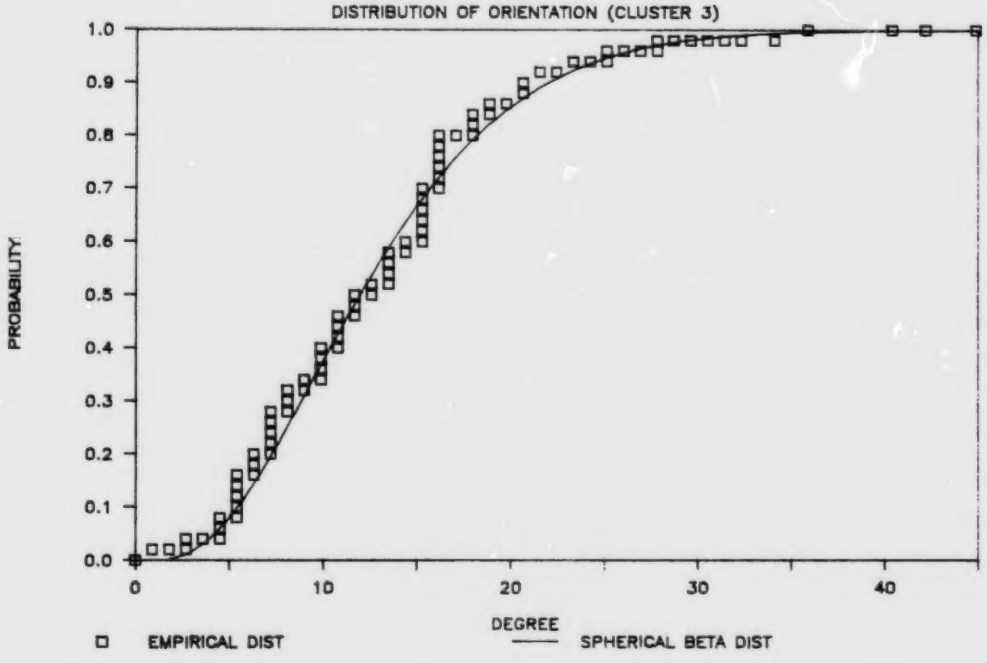
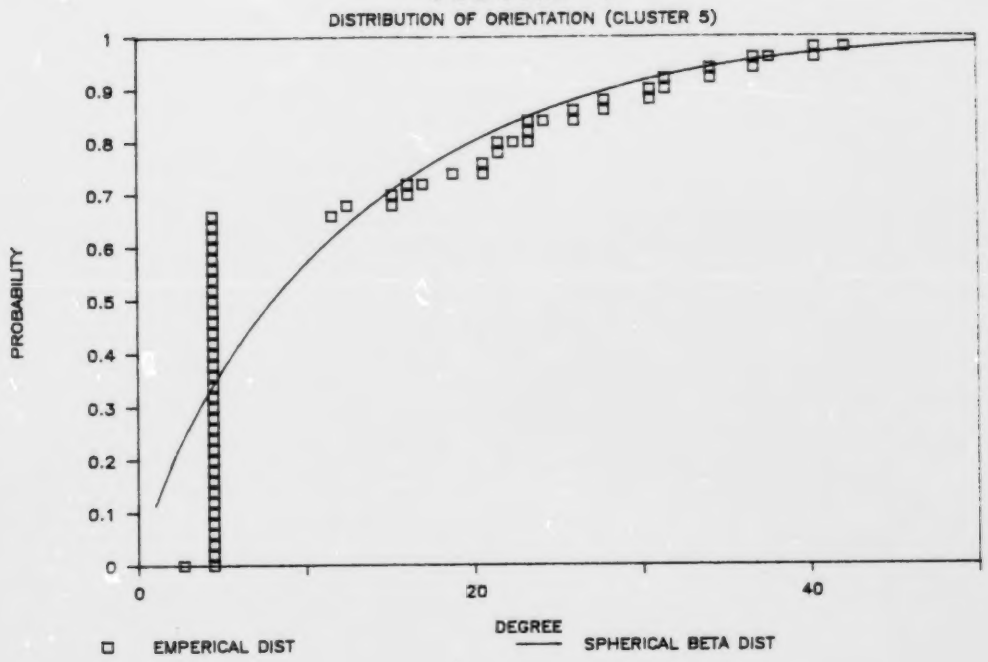


FIGURE 6



3.3 Estimates of the Density Distribution

The possible candidates for the density distribution are the Gamma and the Exponential distributions. After examining the results, it was decided that the Gamma distributions fitted the data better. The following table summarizes the results.

TABLE 4

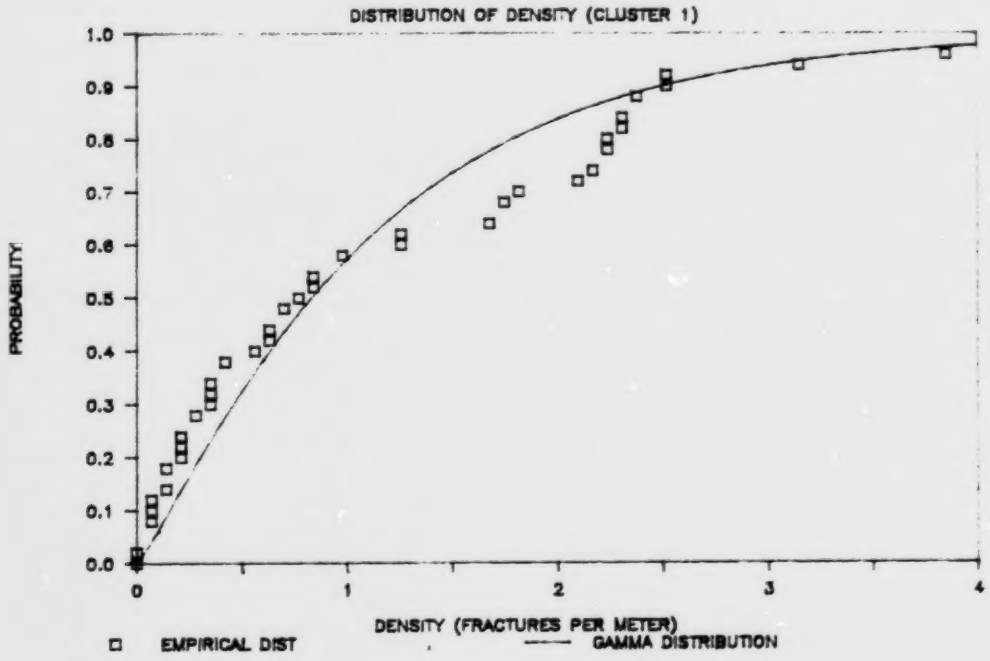
Results of Fitting the Density to the Gamma Distributions

Cluster	α	β	P_o	H_o
1	1.1426	1.0100	0.80	Accepted
2	0.6217	2.0040	0.60	Accepted
3	0.8040	0.5869	0.83	Accepted
4	1.1013	0.4255	0.40	Accepted
5	0.9583	0.1288	0.50	Accepted
6	1.1804	0.4786	0.40	Accepted

where α and β are the estimated parameters of a Gamma distribution and P_o is the probability to reject the null hypothesis H_o.

Remark: Figure 7 shows, as an example, the plot of the empirical distribution of cluster 1 versus the theoretical Gamma distribution. Similar reults for the URL borelhole and excavation data are given in Appendices V and VI, respectively.

FIGURE 7



4. SUMMARY AND DISCUSSIONS

In summary, the following procedures for the estimation of a joint probability distribution were presented and discussed in this report:

- The procedure to determine the joint sets of the fracture orientations (Section 2.3).
- The procedure to estimate the distributions of fracture orientations (Section 2.4).
- A methodology to determine the distributions of fracture dimensions from the surface information (Section 2.5)
- The procedures to estimate the distribution of fracture density (Section 2.6).

The report also provides:

- An example, using the Atikokan data, to demonstrate the above procedures (Section 3); and similar results for the URL borehole data and the excavation data were given in Appendix V and Appendix VI, respectively.

- The computer codes KMNS and BETPLT to carry out the cluster analysis and fitting of the distributions (Appendices I and II).

There are two points that must be reiterated. They are:

- Due to the lack of relevant data and, more importantly, lack of consensus, detailed procedures for neither the estimation of the dimension parameter distributions nor the extrapolation of the joint probability distribution of the fracture parameters to outside the testing area were given.

- The spacing parameter that was used in [2] is equivalent to the density parameter studied here.

The procedures presented in this report performed quite well on the Atikokan data as well as on the URL borehole data. More studies will be needed to investigate the criteria for the extrapolation of the estimated distributions to the entire rock mass.

Approved:

J.G. Cassan
Manager
Operations Research Dept

Submitted:

A.H.C. Chan
Mathematician
Reliability and Statistics Section

A.L. Punhani
Senior Mathematician
Reliability and Statistics Section

AHC/ALP:cmh

REFERENCES

1. Mathers, W.G. and Hembroff, L.R. *Analysis of Flow Through Fracture Rock by the Hele-Shaw Method*, AECL TR-392, in preparation.

2. Chan, A.H., Wilson, L.D. and Punhani, A.L., *Generation of Synthetic Fracture Parameters for Crack Network Analysis*, AECL Technical Report TR-235 (1984).

3. Reilly, J.W., *MAPNET: A Mathematical Model Representing Fracture Network Structure*, Ontario Hydro Research Report 83-219-K (1983).

4. Chung, K.L., *A Course in Probability Theory*, Academic Press, N.Y. ,(1974).

5. Brown, P.A., et al, *The Role of Fracture Characterization in Data Synthesis Modelling*, AECL Technical Report TR-191 (1982).

6. Hartigan, J.A., *Clustering Algorithms*, Wiley, N.Y. (1975)

7. Hartigan, J.A. and Wong, M.A., *K-Mean Clustering Algorithm*, J. of the Royal Statistics Society (Series C), Vol. 28, No. 1 (1979).

8. Miller, A.J., *Selecting Subsets of Regressor Variables*, Proceeding of the Symposium on Regression Analysis (1979), Stat. Society of Australia, NSW Branch.

9. Harr, M., *Mechanics of Particulate Media*, McGraw-Hill, N.Y., (1977).

10. Dienes, J.K. *On the Inference of Crack Statistics from Observations on an Outcroping*, 20th Symposium on Rock Mechanics (1979).

11. Gibbons, J.D., *Non-parametric Statistical Inferences*, McGraw-Hill, N.Y. (1971).

APPENDIX I

SOURCE CODE FOR THE SUBROUTINE KMNS

```
      SUBROUTINE KMNS(A,M1,M,N,C,K1,K,IC1,IC2,NC,AN1,AN2,ANRM,
     *    NCP,D,AK,ITRAN,LIVE,ITER,WSS,IFAULT)
C
C       ALGORITHM AS 136 APPL. STATIST. (1979) VOL.28,NO.1
C
C       DIVIDE M POINTS IN N-DIMENSIONAL SPACE INTO K CLUSTERS
C       SO THAT THE WITHIN CLUSTER SUM OF SQUARES IS MINIMIZED.
C
      DIMENSION A(M1,N),IC1(M),IC2(M),D(M),FLAG(50)
      DIMENSION C(K1,N),NC(K),AN1(K),AN2(K),NCP(K)
      DIMENSION ITRAN(K),LIVE(K),WSS(K),DT(2),ANRM(K)
      DIMENSION AK(K)
C
C       DEFINE BIG TO BE A VERY LARGE POSITIVE NUMBER
C
      BIG=10**10
      IFAULT=3
      IF(K.LE.1.OR.K.GE.M)RETURN
C
C       FOR EACH POINT I, FIND ITS TWO CLOSEST CENTRES,
C       IC1(I) AND IC2(I) .  ASSIGN IT TO IC1(I).
C
      DO 50 I=1,M
      DO 11 IL=1,2
      DT(IL)=0.0
      DA=0.0
      DO 10 J=1,N
      DA=DA+(A(I,J)+C(IL,J))**2.0
      DT(IL)=DT(IL)+(A(I,J)-C(IL,J))**2.0
10    CONTINUE
      FLAG(IL)=1.
      IF(DA.LT.DT(IL))THEN
      FLAG(IL)=-1.
      DT(IL)=DA
      END IF
11    CONTINUE
      IC1(I)=FLAG(1)
      IC2(I)=2*FLAG(2)
      IF(DT(1).LE.DT(2))GO TO 20
      IC1(I)=2*FLAG(2)
      IC2(I)=1*FLAG(1)
      TEMP=DT(1)
      DT(1)=DT(2)
      DT(2)=TEMP
```

```
   20   DO 50 L=3,K
         FLAG(L)=1.
         DB=0.0
         DC=0.0
         DO 30 J=1,M
         DC=DC+(A(I,J)+C(L,J))**2.0
         DB=DB+(A(I,J)-C(L,J))**2.0
         IF(DB.GT.DT(2).AND.DC.GT.DT(2))GO TO 50
   30   CONTINUE
         IF(DC.LT.DB)THEN
         FLAG(L)=-1.
         DB=DC
         END IF
         IF(DB.GE.DT(1))THEN
         DT(2)=DB
         IC2(I)=L*FLAG(L)
         ELSE
         DT(2)=DT(1)
         IC2(I)=IC1(I)
         DT(1)=DB
         IC1(I)=L*FLAG(L)
         END IF
   50   CONTINUE
C
C       UPDATE CLUSTER CENTRES TO BE THE AVERAGE
C       OF POINTS CONTAINED WITHIN THEM
C
         CALL CLSTR(A,M1,N,C,K1,M,K,IC1,NC,1,ANRM,AK)
C
C       CHECK TO SEE IF THERE IS ANY EMPTY CLUSTER AT THIS STAGE
C
         IFAULT=1
         DO 100 L=1,K
         IF(NC(L).EQ.0)RETURN
  100   CONTINUE
         IFAULT=0
         DO 120 L=1,K
         AA=NC(L)
C
C       INITIALIZE AN1,AN2,ITRAN AND NCP
C       AN1(L) IS EQUAL TO NC(L) / (NC(L)-1)
C       AN2(L) IS EQUAL TO NC(L) / (NC(L)+1)
C       ITRAN(L)=1 IF CLUSTER L IS UPDATED IN THE QUICK-TRANSFER STAGE
C       ITRAN(L)=0 OTHERWISE
C       IN THE OPTIMAL-TRANSFER STAGE, NCP(L) INDICATES THE STEP AT
C       WHICH CLUSTER L IS LAST UPDATED
C       IN THE QUICK-TRANSFER STAGE, NCP(L) IS EQUAL TO THE STEP AT
C       WHICH CLUSTER L IS LAST UPDATED PLUS M
C
         AN2(L)=AA/(AA+1.0)
```

```
            AN1(L)=BIG
            IF(AA.GT.1.0)AN1(L)=AA/(AA-1.0)
            ITRAN(L)=1
            NCP(L)=-1
  120     CONTINUE
            INDEX=0
            DO 140 IJ=1,ITER
C
C       IN THES STAGE, THERE IS ONLY ONE PASS THROUGH THE DATA.
C       EACH POINT IS REALLOCATED, IF NECESSARY, TO THE CLUSTER
C       THAT WILL INDUCE THE MAXIMUM REDUCTION IN WITHIN-CLUSTER
C       SUM OF SQUARES.
C
            CALL OPTRA(A,M1,M,N,C,K1,K,IC1,IC2,NC,AN1,AN2,NCP,
     *      D,ITRAN,LIVE,INDEX)
C
C       STOP IF NO TRANSFER TOOK PLACE IN THE LAST M
C       OPTIMAL-TRANSFER STEPS
C
            IF(INDEX.EQ.M)GO TO 150
C
C       EACH POINT IS TESTED IN TURN TO SEE IF IT SHOULD BE
C       REALLOCATED TO THE CLUSTER WHICH IT IS MOST LIKELY TO
C       BE TRANSFERRED TO (IC2(I)) FROM ITS PRESENT CLUSTER(IC1(I)).
C       LOOP THROUGH THE DATA UNTIL NO FURTHER CHANGE IS TO TAKE
C       PLACE.
C
            CALL QTRAN(A,M1,M,N,C,K1,K,IC1,IC2,NC,AN1,AN2,
     *    NCP,D,ITRAN,INDEX)
C
C       IF THERE ARE ONLY TWO CLUSTERS,
C       NO NEED TO RE-ENTER OPTIMAL-TRANSFER STAGE
C
            IF(K.EQ.2)GO TO 150
C
C       NCP HAS TO BE SET TO 0 BEFORE ENTERING OPTRA
C
            DO 130 L=1,K
  130     NCP(L)=0
  140     CONTINUE
C
C       SINCE THE SPECIFIED NUMBER OF INTERACTIONS IS EXCEEDED
C       IFAULT IS SET TO BE EQUAL TO 2.
C       THIS MAY INDICATE UNFORESEEN LOOPING.
C
            IFAULT=2
C
C       COMPUTE WITHIN CLUSTER SUM OF SQUARES FOR EACH CLUSTER.
C
  150     CALL ZAP(WSS,K)
```

```
      CALL CLSTR(A,M1,N,C,K1,M,K,IC1,NC,0,ANRM,AK)
      DO 210 I=1,M
      WS=0.0
      II=ABS(IC1(I))
      DUM=1.
      IF(II.NE.IC1(I))DUM=-1.
      DO 200 J=1,N
      DA=DUM*A(I,J)-C(II,J)
      WS=WS+DA*DA
200   CONTINUE
      WSS(II)=WSS(II)+WS
210   CONTINUE
      RETURN
      END
```

APPENDIX II

SOURCE CODE FOR THE SUBROUTINE BETPLT

```
C
C   SUBROUTINE BETPLT
C
C   SUBROUTINE BETA PLOT.
C  THIS SUBROUTINE CALCULATES MEAN AND STANDARD DEVIATION
C  OF THE VECTOR 'VECT' AND THEN GIVEN THE MAXIMUM AND
C  MINIMUM VALUES OF THE FUNCTION IT CALCULATES THE ACTUAL
C  AND THEORETICAL CUMULATIVE DISTRIBUTION FUNCTIONS. IT
C  ALSO PLOTS THE TWO CDF'S.
C
C  ***INPUT PARAMETERS:
C
C  VECT: THE VECTOR CONTAINING THE DATA VALUES.
C
C  N    : THE SIZE OF VECT.
C
C  A    : THE MINIMUM VALUE POSSIBLE FOR THE DATA.
C
C  B    : THE MAXIMUM VALUE POSSIBLE FOR THE DATA.
C
C  TIN : THE INPUT UNIT ASSOCIATED WITH THE FILE CONTAINING
C        THE TITLES FOR THE PLOT.  THE FILE MUST BE OF THE
C        FORM -LINE #1- THE MAIN TITLE CENTERED IN COLUMNS
C                        1-72.
C             -LINE #2- THE XTITLE CENTERED IN COLUMNS 1-36.
C             -LINE #3- THE YTITLE CENTERED IN COLUMNS 1-36.
C        ***NOTE IF TIN.LE.0 THEN NO TITLES ARE USED.
C
C  OPN: OPTIONS, WHERE
C          OPN=0 BETA DISTRIBUTION
C          OPN=1 GAMMA
C          OPN=2 EXPONENTIAL
C
C  ***OUTPUT PARAMETERS:
C
C  MEAN: THE MEAN VALUE OF VECT.
C
C  SD   : THE STANDARD DEVIATION OF VECT.
C
C  ALPHA:THE ALPHA VALUE FOR THE BETA FUNCTION.
C
C  BETA: THE BETA VLAUE FOR THE BETA FUNCTION.
C
C  KSMR: A TWO ELEMENT VECTOR CONTAINING THE PROBABILITIES FOR A
```

```
C         ONE-WAY AND TWO-WAY KOLMOGOROV-SMIRNOV TEST.
C
C  ***OTHER ROUTINES USED:
C
C  VSRTA:   AN IMSL ROUTINE WHICH SORTS A VECTOR INTO ASCENDING
C           ORDER.
C
C  MDBETA:  AN IMSL ROUTINE USED TO CALCULATE THE BETA PROBABILITY
C
C  MDSMR:   AN IMSL ROUTINE USED TO CALCULATE THE KOLMOGOROV-
C           SMIRNOV STATISTICS.
C
C  USPLO:   AN IMSL ROUTINE USED TO PLOT THE CDF FUNCTIONS
C
             SUBROUTINE
             BETPLT(VECT,N,A,B,TIN,OPN,MEAN,SD,ALPHA,BETA,KSMR)
             IMPLICIT LOGICAL(A-Z)
             INTEGER N
             REAL VECT(2000)
             REAL A
             REAL B
             INTEGER TIN
             INTEGER OPN
             REAL MEAN
             REAL SD
             REAL ALPHA
             REAL BETA
             REAL KSMR(2)
             REAL SUM
             REAL SUMSQ
             REAL XT
             REAL VT
             REAL X
             REAL Y(2000,2)
             REAL XR(2000)
             REAL RANGE(4)
             CHARACTER*2 ICHAR
             REAL D
             INTEGER IER
             INTEGER I
             DATA ICHAR/'AT'/
             TIN=0
             D=0
C
C       CHECK TO SEE THAT N IS GREATER THAN TWO.
C
             IF(N.LE.2)THEN
             WRITE(6,101)N
             STOP
             END IF
```

```
C
C     CHECK TO SEE THAT N.GE.100. IF NOT ISSUE WARNING.
C
            IF(N.LT.100)WRITE(6,106)
C
C     CHECK TO SEE THAT A IS LESS THAN B
C
            IF(A.GE.B)THEN
            WRITE(6,102)A,B
            STOP
            END IF
C
C    SORT THE INPUT VECTOR.
C
      3     CALL VSRTA(VECT,N)
C
C    CHECK TO SEE THAT VECT(1).GE.A. IF NO THEN
C    SET A=VECT(1). AND ISSUE A WARNING.
C
            IF(VECT(1).LT.A)THEN
            WRITE(6,104)VECT(1),A
            A=VECT(1)
            END IF
C
C    CHECK TO SEE THAT VECT(N).LE.B . IF NOT THEN
C    SET B=VECT(N), AND ISSUE A WARNING.
C
            IF(VECT(N).GT.B)THEN
            WRITE(6,105)VECT(N),B
            B=VECT(N)
            END IF
C
C     CALCULATE SUM OF SQUARES AND SUM OF VECT.
C
            SUM=0
            SUMSQ=0
            DO 1 I=1,N,2
            SUM=SUM+VECT(I)
      1   CONTINUE
C
C     CALCULATE MEAN & SD & ALPHA & BETA.
C
            MEAN=SUM/(N/2)
            DO 11 I=1,N,2
            SUMSQ=SUMSQ+(VECT(I)-MEAN)**2
      11  CONTINUE
            SD=SQRT(SUMSQ/((N/2)-1))
            IF(OPN.EQ.0)THEN
            XT=(MEAN-A)/(B-A)
            VT=(SD/(B-A))**2
```

```
            ALPHA=XT**2/VT*(1-XT)-XT
            BETA=(ALPHA)/XT-ALPHA
            END IF
            IF(OPN.EQ.1)THEN
            BETA=(SD**2)/MEAN
            ALPHA=(MEAN/BETA)-1.
            VT=ALPHA+1.
            END IF
C
C     CALCULATE THE FUNCTION VALUES.
C
            DO 2 I=1,N
            Y(I,1)=I/FLOAT(N)
            IF(OPN.EQ.0)THEN
            X=(VECT(I)-A)/(B-A)
            CALL MDBETA(X,ALPHA,BETA,Y(I,2),IER)
            END IF
            IF(OPN.EQ.1)THEN
            X=VECT(I)/BETA
            CALL MDGAM(X,VT,Y(I,2),IER)
            END IF
            IF(OPN.EQ.2)Y(I,2)=1.0-EXP((A-VECT(I))/SD)
            IF(I.NE.1)D=AMAX1(D,AMAX1(ABS(Y(I,1)-Y(I,2)),
           *                ABS(Y(I,1)-Y(I-1,2))))
         2  CONTINUE
            D=SQRT(N)*D
            WRITE(6,107)D,N
       107  FORMAT(10(/),9X,'THE KOLMOGOROV-SMIRNOV STAT =',F14.8,
           *     //,9X,'THE NUMBER OF OBSERVATIONS =',I10/)
C
C     CALCULATE THE KOLMOGOROV-SMIRNOV STATISTICS.
C
            CALL MDSMR(D,KSMR(1),KSMR(2))
C
C     SET UP THE PLOTTING PARAMETERS.
C
            RANGE(1)=A
            RANGE(2)=B
            RANGE(3)=0.
            RANGE(4)=1.
C
C     PLOT THE ACTUAL AND THEORETICAL CDF'S.
C
            CALL USPLO(VECT,Y,2000,N,2,1,'VECT VERSUS Y',13,'VECT',
           *           4,'Y',1,RANGE,ICHAR,1,IER)
         8  CONTINUE
C
C     NOW PLOT THE TWO CDF'S VS EACH OTHER.
C
            DO 7 I=1,N
```

```
      XR(I)=Y(I,2)
7     CONTINUE
      RANGE(1)=0
      RANGE(2)=1
      CALL USPLO(XR,Y,2000,N,2,1,'XR VERSUS Y',11,'XR'
     *           ,2,'Y',1,RANGE,ICHAR,1,IER)
      RETURN
98    WRITE(6,103)
      GO TO 8
99    WRITE(6,103)
      GO TO 3
100   FORMAT(72A1/36A1/36A1)
101   FORMAT('1***ERROR IN BETPLT: N IS LESS THAN 2. N= ',I4)
102   FORMAT('1***ERROR IN BETPLT: A.GE.B. A= ',F8.2,' B= ',
     *       F8.2)
103   FORMAT('0***ERROR IN BETPLT: END OF FILE ENCOUNTERED IN '
     *      ,'PLOT TITLES FILE.'/' NO TITLES WILL BE USED.')
104   FORMAT('0***WARNING IN BETPLT: A.GT.VECT(1). A= ',F8.2,
     *      ' VECT(1)= ',F8.2/' A WILL BE SET EQUAL TO VECT(1).')
105   FORMAT('0***WARNING IN BETPLT: B.LT.VECT(N), B= ',F8.2,
     *      ' VECT(N)= ',F8.2/' B WILL BE SET EQUAL TO VECT(N).')
106   FORMAT('0***WARNING IN BETPLT: THE KSMR VALUES WILL BE',
     *      ' INACCURRATE SINCE N IS LESS THAN 100.')
      END
```

APPENDIX III

THE KOLMOGOROV-SMIRNOV-DURBIN STATISTICS

The mathematical foundation for the BETPLT subroutine is based on the Kolmogorov-Smirnov-Durbin statistic for the goodness-of-fit test. Let $x_1, x_2, \ldots, x_N$ be the order samples, ie $x_1 < x_2 < \ldots < x_N$ from a certain population. Suppose that one wants to test if the samples are drawn from a population having a probability distribution $F(x;\Theta)$, where $\Theta = (\theta_1, \ldots, \theta_k)$ is the vector of parameters of F (eg mean, standard deviation, ..., etc). If Θ is known, the Kolmogorov-Smirnov (K-S) statistic is defined by

$$D_N = \sqrt{N} \max | F_N(x) - F(x;\Theta) | \qquad \text{(III.1)}$$

where $F_N(x)$, the empirical distribution constructed from the samples, can be used to test the null hypothesis. The computational procedures for D_N can be found in [10].

In general, however, these parameters are not known. So, intuitively, substituting the estimated parameters Θ_n for Θ in equation (III.1) may provide a similar measure of the goodness-of-fit, ie

$$D_N^* = \sqrt{N} \max | F_N(x) - F(x;\Theta_n) | \qquad \text{(III.2)}$$

This turns out to be almost corrected. The only problem is that the statistic D_N^* does not have a tractable distribution if Θ_n is estimated from all the samples. Hence, its usefulness is limited.

Fortunately, Durbin proved that if $\Theta_{n/2}$ is an estimate of Θ using only half (randomly selected) of the samples, then substituting $\Theta_{n/2}$ for Θ in equation (III.1) will result in a statistic with the same distribution as D_N. This statistic, namely,

$$D_N^{**} = \sqrt{N} \max | F_N(x) - F(x;\Theta_{n/2}) | \qquad \text{(III.3)}$$

is called the Durbin-Kolmogorov-Smirnov (D-K-S) test.

Now, to estimate the distributions of fracture parameters, it is simply to test the goodness-of-fit of the data against all possible distributions. Clearly, this is impossible. Therefore, a few plausible robust distributions were selected to be tested. The Beta, Exponential and Gamma distributions were selected in this task.

The outline of the BETPLT subroutine are as follows:

1. Input the data $x_1, x_2, \ldots, x_N$ and the distribution to be fitted;
2. Sort the data x_i's such that $x_1 < x_2 < \ldots < x_N$;
3. Select a random subsample from the x_i's;
4. Compute $\Theta_{n/2}$ from the subsample of the selected distribution.

Let K=[N/2]. Compute

$$m_x = \Sigma_i\, x_i/K\,, \quad s^2 = \Sigma_i\,(x_i - m_x)^2/(K\text{-}1)$$

$$\min = \min\{x_1, x_2, \ldots, x_K\},\ \max = \max\{x_1, x_2, \ldots, x_K\}$$

$$M_x = [m_x - \min]/[\max - \min]$$

Let $\Theta_{n/2} = (\alpha,\beta)$. Then, for the Beta distribution, compute

$$\alpha = [M_x(1 - M_x)/2] - M_x$$

$$\beta = [\alpha/M_x] - \alpha$$

For the Gamma distribution, compute

$$\alpha = m_x^2/s^2$$

$$\beta = m_x/s^2$$

5. Compute $F(x_i;\Theta_{n/2})$ and $h_i = |\,(i/N) - F(x_i;\Theta_{n/2})\,|$;
6. Determine the maximum of h_i's, say h_{max}, and define

$$D_N^{**} = \sqrt{N}\, h_{max};$$

7. Determine the probability of D_N^{**};
8. Plot
   (a) i/N and $F(x_i;\Theta_{n/2})$ versus x_i;
   (b) i/N versus $F(x_i;\Theta_{n/2})$.

APPENDIX IV

ESTIMATION OF FRACTURE DIMENSIONS DISTRIBUTION FROM THE DISTRIBUTION OF SURFACE FRACTURE LENGTHS

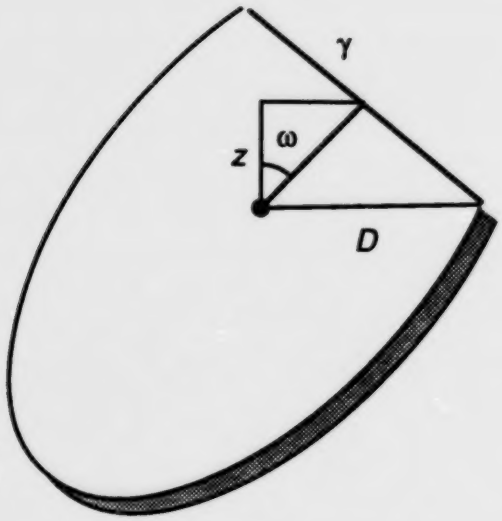
Since fracture dimensions are not observable, it would be impossible to estimate its distribution. Fortunately, if a fracture is idealized as a circular disk, the diameter of a fracture can be deduced from the trace where it intersecting the surface. There is however, no data on the fracture aperture. At this point, it can be assume to be independent from all other parameters. This way, without lost of generality, assume it be to constant. Hence, the distribution needed to be estimated is that of the diameters.

Consider a fracture intersects the surface whose dip angle is ω. Let D denote its diameter and L the length of its trace on the surface. From Figure 8, the probability of $L > \gamma$, denoted by $\Psi(\gamma)$, can be expressed by the equation

$$\Psi(\gamma) = \int_{\gamma/2}^{\infty} \int_{\Omega} P\{ |Z| < \sin\omega \sqrt{(\zeta^2 - (\gamma/2)^2)} \mid D = \zeta \} dF(\zeta) dH(\omega)$$

FIGURE 8

RELATIONSHIP BETWEEN SURFACE LENGTH AND FRACTURE DIAMETER



$$z = \sin\omega \sqrt{D^2 - (\gamma/2)^2}$$

where F and H are the probability distribution functions of the of diameters and of the dip angle ω, respectively. Firstly, the dimensions, location and orientation parameters are assumed to be independent. Secondly, suppose that there are N_o fractures in the host rock mass whose depth is Z_o. The number fractures whose centre fall above a depth of z, say n_o, is proportional to Z_o. In other words, $n_o/N_o = z/Z_o$. This implies that

$$\int_\Omega P\{ |Z| < \sin\omega \, [(\zeta^2 - (\gamma/2)^2)^{1/2} \} dH(\omega) = C_o \, [(\zeta^2 - (\gamma/2)^2)]^{1/2}$$

where $C_o = [\int_\Omega \sin\omega \, dH(\omega)]/Z_o$ and $\Omega = [0,\pi/2]$.

This infers that the distribution function F can be determined by solving the integral equation:

$$\Psi(\gamma) = C_o \int_{\gamma/2}^{\infty} [(\zeta^2 - (\gamma/2)^2)]^{1/2} \, F'(\zeta) \, d\zeta \qquad \text{(IV.1)}$$

where F' is the derivative of F.

Substituting x for $(\gamma/2)^2$ and ξ for ζ^2, the above equation can be expressed by

$$\Psi^*(x) = C_o \int_x^{\infty} (\xi - x)^{1/2} \, f(\xi) \, d\xi$$

Differentiate both side with respect to x, we get

$$\psi(x) = C_o \int_x^{\infty} (\xi - x)^{-1/2} \, f(\xi) \, d\xi \qquad \text{(IV.2)}$$

where $\Psi^*(x) = \Psi(2\sqrt{x})$, $\psi(x) = d\Psi^*(x)/dx$ and $f(\xi) = F'(\sqrt{\xi})/2\sqrt{\xi}$. Then, problem is simply to solve the integral equation (IV.2).

To solve this integral equation, the following result is needed:

<u>Theorem</u> The solution for the integral equation

$$u(x) = \int_x^{\infty} (\xi - x)^{-1/2} \, f(\xi) d\xi$$

is

$$\int_z^{\infty} f(x) dx = \pi^{-1} \int_z^{\infty} (x - z)^{-1/2} \, u(x) dx$$

or

$$f(z) = \pi^{-1} \, D_z \int_z^{\infty} (z - x)^{-1/2} \, u(x) dx$$

where D_z denotes the differentiation with respect to z.

The following lemmas are needed in the proof of the theorem. They will be stated without proofs.

Lemma 1 If $0 < \lambda < 1$, then

$$\int_z^{\xi}(x-z)^{\lambda-1}(\xi-x)^{-\lambda}\,dx = \pi/\sin\lambda\pi$$

Lemma 2 If $0 < \lambda < 1$ and $u(\xi)$ differentiable and bounded, then

$$\int_z^{\infty}[\int_z^{\xi}(x-z)^{\lambda-1}(\xi-x)^{-\lambda}\,u(\xi)dx\,]d\xi = \int_z^{\infty}[\int_x^{\infty}(x-z)^{\lambda-1}(\xi-x)^{-\lambda}\,u(\xi)d\xi\,]dx$$

Now, from Lemma 1 with $\lambda = 1/2$, multiple both side by $f(\xi)$ and integrate with respect to ξ, we get

$$\pi\int_z^{\infty}f(\xi)d\xi$$

$$= \int_z^{\infty}[\int_z^{\xi}(z-x)^{-1/2}(x-\xi)^{-1/2}\,f(\xi)dx]d\xi$$

$$= \int_z^{\infty}[\int_x^{\infty}(z-x)^{-1/2}(x-\xi)^{-1/2}\,f(\xi)d\xi\,]dx$$

$$= \int_z^{\infty}[(z-x)^{-1/2}u(x)]dx$$

This proves the theorem.

Now, from the above theorem, the distribution function F is given by

$$F(y) = 1 - C\int_{2y}^{\infty}[(\gamma/2)^2-y^2]^{-1/2}\,d\Psi(\gamma)$$

where $C = 1/\pi C_o$, $C_o = [\int_\Omega \sin\omega\ dH(\omega)]/Z_o$, Z_o is the depth of the rock mass and $\Omega = [0,\pi/2]$.

The rationale of this result is based on that of [10]. However, there is slight oversight in a solution given here. The solution given here is also applicable to the solution in [10].

APPENDIX V

RESULTS FROM THE URL BOREHOLE DATA

The Geological Survey of Canada (GSC) provided Ontario Hydro with the URL borehole data. The format of the data, which is stored as a FORTRAN file, is as follows:

Each record consists of eight fields pertaining to a fracture. They are
(BN,DE,FT,SK,DP,TF,WE,FF)

with format (I1,F9.3,A2,I3,I2,F6.1,F4.1,5A2)

where
FT = fracture type,
SK = strike angle,
DP = dip angle,
TF = thickness of the fracture,
FF = fracture filling type,
BN = borehole number,
DE = depth from surface, and
WE = weighting factor.

The data that are not usable, are deleted as did in Section 3. Unlike the Atikokan data, the coordinates of the fractures are not available. To determine the density of the fractures, the following formula is used:

Let $f_1,...,f_n$ be the fractures fell between the depths of mk and $m(k-1)$ for some k, where m is the fixed length. Let DE_i and WE_i denote the DE and WE of f_i. The density observed from depth between mk and $m(k-1)$ is

$$p = (DE_1WE_1 + ... + DE_nWE_n)/m.$$

The results are outline as following:

Determination of the Joint Sets

A series of K-Mean cluster analyses were performed for $K = 3$ to $K = 8$. From Figure 9, it was apparent that the reduction of total sum of squares is almost linear to the number of clusters. Therefore, from the total sum of square reduction point of views, the optimal number of clusters can not be determine subjectively. On the other hand, when the sum of squares per cluster were plotted against N (see Figure 10), it is easy to select $N = 6$ as the optimal number.The following tables summarize the results of the cluster analysis. Table V.1 provides the average orientation of each of the clusters.

FIGURE 9

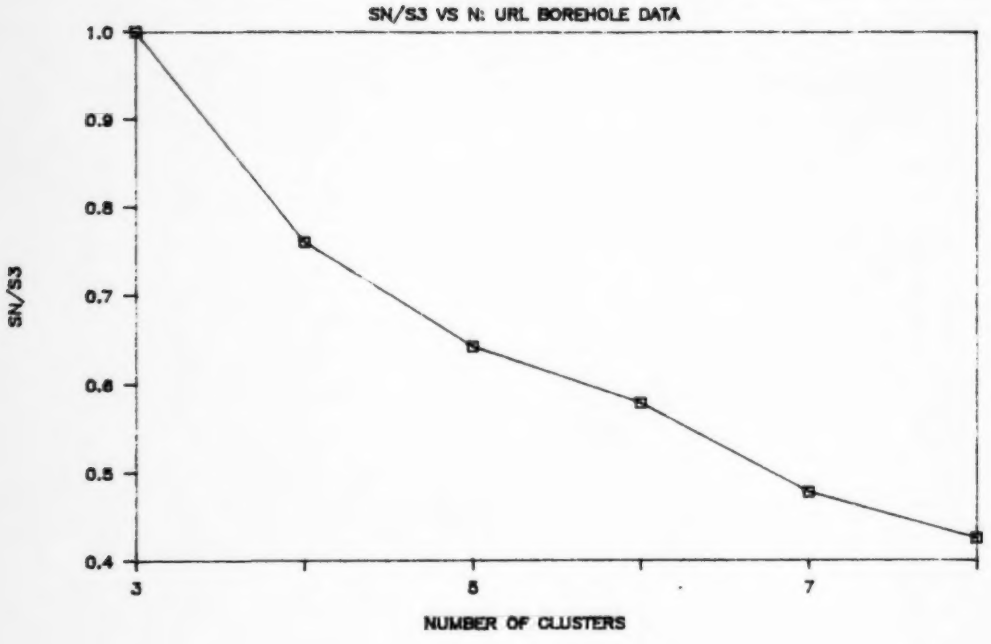


FIGURE 10

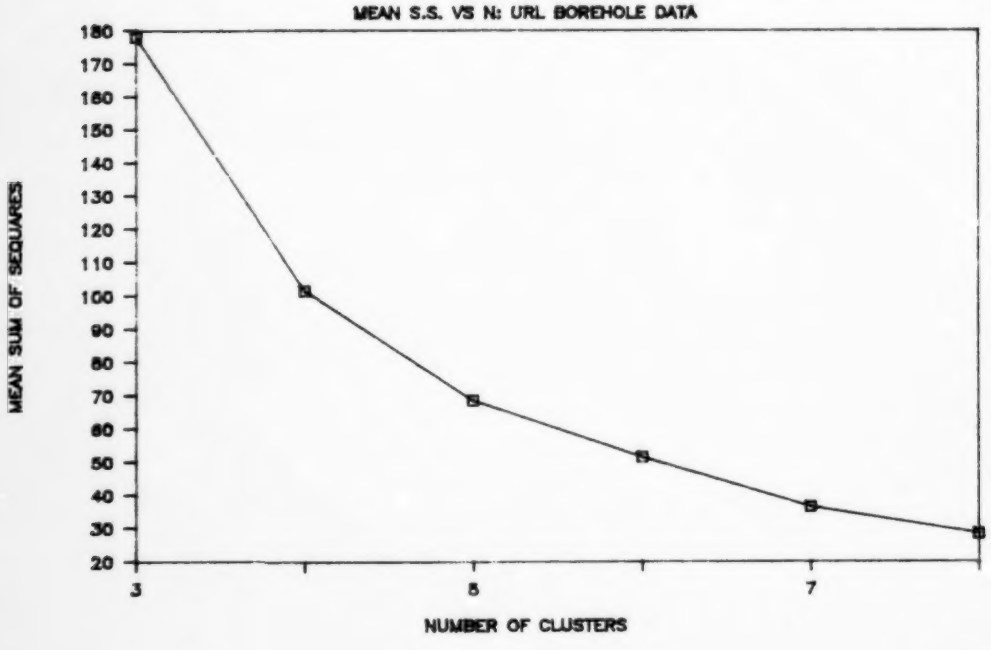


TABLE V.1

Cluster Analysis Results on the Orientation

Cluster	Center of Orntns Strike	Dip	Within Clsrs Sum of Squrs	No. of Obsvns
1	94.9	11.4	38.1	579
2	93.6	23.6	48.2	486
3	173.9	25.5	36.9	463
4	12.7	24.4	34.4	450
5	135.8	86.5	85.0	379
6	29.4	81.7	66.5	271

Remark: The above results were obtained from the original data file provided by GSC which does not include the value AV. This is rather important because the AV value will give more weight to the fractures as they are almost parallel to the angles of the boreholes.

Estimates of the Orientation Distributions

From the cluster analysis, the cluster numbers JS are added to the data file. So, from the data file, all the fractures from a cluster (joint set) can be selected. As described in Section 2.3, the Spherical Beta distribution is probably the most robust one for the data. The results are summarized in Table V.2.

TABLE V.2

Results of Fitting the Orientation to the Spherical Beta Distributions

Cluster	α	β	P_o	H_o
1	2.64	2.04	0.39	Accepted
2	2.53	2.10	0.36	Accepted
3	1.97	1.65	0.94	Accepted
4	2.58	2.09	1.00	Rejected
5	2.41	2.17	0.99	Rejected
6	2.06	1.88	0.98	Accepted

where α and β are the estimated parameters of the Spherical Beta distribution (see Definition in 2.4.1); H_o is the null hypothesis that the orientation of a cluster have a Spherical Beta distribution; P_o is the probability to reject H_o.

Although clusters 4 and 5 were rejected, for the purpose of simulation, they may also be considered to have spherical distributions.

Estimates of the Density Distribution

The possible candidates for the density distribution are the Gamma and the Exponential distributions. After examining the results, it was decided that the Gamma distributions fitted the data better. The following table summarizes the results.

TABLE V.3

Results of Fitting the Density to the Gamma Distributions

Cluster	α	β	P_o	H_o
1	0.49	2.64	0.75	Accepted
2	0.43	2.29	0.92	Accepted
3	0.81	2.17	0.74	Accepted
4	0.91	1.73	0.69	Accepted
5	0.87	2.05	0.38	Accepted
6	1.01	1.94	0.55	Accepted

where α and β are the estimated parameters of a Gamma distribution and P_o is the probability to reject the H_o

APPENDIX VI

RESULTS FROM THE URL EXCAVATION DATA

The Atomic Energy of Canada Ltd. (AECL) provided Ontario Hydro with the URL excavation data. The format of the data, which is stored as a FORTRAN file, is as follows:

Each record consists of five fields pertaining to a fracture. They are

(BN,DE,FT,SK,DP)

with format (10X,A1,I3,19X,A1,3X,I3,I2)

where

FT = fracture type,
SK = strike angle,
DP = dip angle,
BN = borehole number, and
DE = depth from surface.

The data that are not usable, are deleted as did in Section 3. Note that the weighting factor WE provided in the URL borehole data is missing here. Therefore, the weighting factor for each fracture is assumed to be 1 and the distributions of the density parameter for the joint sets may be biased. The results are outline as following.

The number of joint sets are decided to be 5 (See Figure 11). The following tables summarize the results of the cluster analysis. Table VI.1 provides the average orientation of each of the clusters.

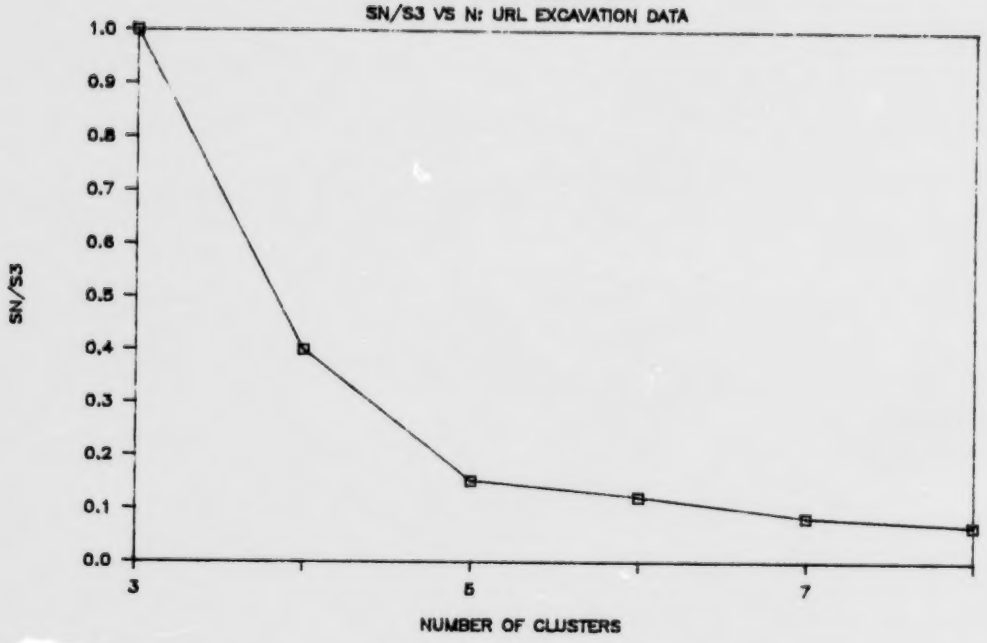
TABLE VI.1

Cluster Analysis Results on the Orientation

Cluster	Center of Orntns Strike	Dip	Within Clsrs Sum of Squrs	No. of Obsvns
1	149.1	7.3	21.5	272
2	301.1	86.9	15.1	980
3	287.8	90.4	22.1	1214
4	145.1	89.5	10.2	174
5	52.9	103.5	16.2	334

Remark: The above results were obtained from the original data file provided by GSC which does not include the value AV. This is rather important because the AV value will give more weight to the fractures as they are almost parallel to the angles of the boreholes.

FIGURE 11



Estimates of the Orientation Distributions

From the cluster analysis, the cluster numbers JS are added to the data file. So, from the data file, all the fractures from a cluster (joint set) can be selected. As described in Section 2.3, the Spherical Beta distribution is probably the most robust one for the data. The results are summarized in Table VI.2.

TABLE VI.2

Results of Fitting the Orientation to the Spherical Beta Distributions

Cluster	α	β	P_o	H_o
1	1.00	2.57	0.89	Accepted
2	1.00	4.63	1.00	Rejected
3	1.00	6.12	1.00	Rejected
4	1.00	3.07	0.96	Accepted
5	1.00	3.69	0.98	Accepted

where α and β are the estimated parameters of the Spherical Beta distribution (see Definition in 2.4.1); H_o is the null hypothesis that the orientation of a cluster have a Spherical Beta distribution; P_o is the probability to reject H_o.

Although clusters 2 and 3 were rejected, for the purpose of simulation, they may also be considered to have spherical distributions.

Estimates of the Density Distribution

The possible candidates for the density distribution are the Gamma and the Exponential distributions. After examining the results, it was decided that the Gamma distributions fitted the data better. The following table summarizes the results.

TABLE VI.3

Results of Fitting the Density to the Gamma Distributions

Cluster	α	β	P_o	H_o
1	0.33	4.51	0.97	Accepted
2	1.11	2.27	0.42	Accepted
3	1.06	3.08	0.75	Accepted
4	0.65	2.07	0.85	Accepted
5	1.75	1.40	0.91	Accepted

where α and β are the estimated parameters of a Gamma distribution and P_o is the probability to reject the null hypothesis H_o.